## Apratoxin E, a Cytotoxic Peptolide from a Guamanian Collection of the Marine Cyanobacterium Lyngbya bouillonii

Susan Matthew,<sup>†</sup> Peter J. Schupp,<sup>‡</sup> and Hendrik Luesch\*,<sup>†</sup>

Department of Medicinal Chemistry, University of Florida, 1600 SW Archer Road, Gainesville, Florida 32610, and University of Guam Marine Laboratory, UOG Station, Mangilao, Guam 96923

Received December 13, 2007

A collection of the marine cyanobacterium *Lyngbya bouillonii* from Guam afforded apratoxin E (1), a new peptide–polyketide hybrid of the apratoxin class of cytotoxins. The planar structure of 1 was elucidated by NMR spectroscopic analysis and mass spectrometry. Configurational assignments of stereocenters in the peptide portion were made by chiral HPLC analysis of the acid hydrolysate. The relative configuration in the polyketide moiety was assigned by comparison of NMR data including proton–proton coupling constants with those of the known analogues. Apratoxin E (1) displayed strong cytotoxicity against several cancer cell lines derived from colon, cervix, and bone, ranging from 21 to 72 nM, suggesting that the  $\alpha$ , $\beta$ -unsaturation of the modified cysteine residue is not essential for apratoxin activity. The 5- to 15-fold reduced activity compared with apratoxin A (2) is attributed to the dehydration in the long-chain polyketide unit, which could affect the conformation of the molecule.

Cyanobacteria produce an array of bioactive secondary metabolites with characteristic biosynthetic signatures, which include a high degree of modified amino acids in a peptide-polyketide hybrid backbone.<sup>1</sup> Representative examples are the apratoxins, which are of mixed biosynthetic origin and exhibit potent cancer cell growth inhibitory activity by inducing G1 phase specific cell cycle arrest and apoptosis.<sup>2</sup> Apratoxins A-C (2-4) were isolated from the remarkably prolific Lyngbya bouillonii collected in Guam and Palau<sup>3,4</sup> and apratoxin D from a variety of the same species collected in Papua New Guinea.<sup>5</sup> The apratoxins have become attractive targets for synthetic efforts, culminating in several total syntheses.<sup>6–8</sup> In our search for natural apratoxin analogues aimed at expanding our knowledge about structure-activity relationships within this intriguing compound class, we have encountered a new natural apratoxin from a variety of L. bouillonii from Guam. Here, we report the isolation, structure determination, and evaluation of the antiproliferative activity of the new apratoxin designated as apratoxin E (1), which was the predominant apratoxin in L. bouillonii inhabiting deeper water and which was not present in the apratoxin A (2) producer that lives in shallower water at the same collection site (Figure 1).

To determine if the structural diversity of apratoxins may also be a function of collection depth, we carried out collections of the apratoxin-producing variety of *L. bouillonii* from shallow depths of 1-3 m and deeper waters of 12-14 m at Finger's Reef, Apra Harbor, Guam. The freeze-dried cyanobacterial collections were extracted with a 1:1 solvent mixture of MeOH and EtOAc and crude extracts fractionated by normal-phase chromatography. The most cytotoxic fraction was subjected to reversed-phase HPLC to afford apratoxin E (1) as an amorphous, colorless solid, along with only trace amounts of apratoxin A (2) in the deep collection, and only apratoxin A (2) from the shallow collection. Compound 1 and lyngbyastatin 2<sup>9</sup> accounted for the majority of the extract's growthinhibitory activity (Figure 1).

The molecular formula of **1** was established as  $C_{43}H_{65}N_5O_7S$ , on the basis of an  $[M + H]^+$  peak in the HRESI/APCIMS at *m/z* 796.4683 (calcd for  $C_{43}H_{66}N_5O_7S$ , 796.4683). <sup>1</sup>H NMR data in CDCl<sub>3</sub> showed the presence of an amide NH and two *N*-Me groups indicative of a peptide, with characteristic chemical shifts remi-



**Figure 1.** Comparison of HPLC profiles derived from two varieties of *L. bouillonii* from Finger's Reef, Apra Harbor (Guam), encountered at different depths. The shallow *L. bouillonii* was collected at 1-3 m depth and the deep *L. bouillonii* at 12-14 m depth. Extracts were separated by silica gel chromatography and the cytotoxic fractions subjected to reversed-phase HPLC (see Experimental Section). Apratoxin E (1) is unique to the deep *L. bouillonii* (red trace) and also contained higher amounts of lyngbyastatin 2 compared with the shallow *L. bouillonii* (blue trace).

niscent of those of apratoxin A (2) and the presence of conformers in a 3:2 ratio. Detailed <sup>1</sup>H and <sup>13</sup>C NMR analysis of **1** in conjunction with COSY, TOCSY, ROESY, HSOC, and HMBC data revealed a hybrid structure comprised of peptide (O-Me-Tyr, N-Me-Ile, N-Me-Ala, and Pro) and polyketide fragments (C27-C31 and C33-C45) that were similar to apratoxin A  $(2)^3$  (Table 1), yet exhibited a lower degree of C-methylation and a different position of the double bond.<sup>10</sup> The apparent disparity between 1 and 2 was the replacement of the olefinic methine and allylic methyl proton signals of the  $\alpha$ . $\beta$ -unsaturated modified cysteine (moCys) spin system (C27–C32) ( $\delta_{H-29}$  6.35,  $\delta_{H-30}$  5.25, and  $\delta_{H-32}$  1.96) by a pair of methylene signals H<sub>2</sub>-28 and H<sub>2</sub>-29 resonating at  $\delta_{\rm H}$  2.76/ 2.49 and 2.21/1.67 (major conformer), respectively.<sup>10</sup> Thus, in compound 1, this unit is neither  $\alpha,\beta$ -unsaturated nor methylated at C28, in contrast to apratoxin A (2). NMR analysis further suggested that the substituted C<sub>9</sub> carboxylic acid unit (C33-C45) more closely resembled the semisynthetic (E)-dehydroapratoxin A (5);<sup>4</sup> both 1 and 5 lack the hydroxyl group at C35 present in 2 and possess a double bond between C34 and C35.<sup>10</sup> The allylic methyl group (C44) was missing, so that compound 1 consisted of a 7-hydroxy-5,8,8-trimethyl-2-nonenoic acid residue (C33-C45). The geometry

<sup>\*</sup> To whom correspondence should be addressed. Tel: (352) 273-7738. Fax: (352) 273-7741. E-mail: luesch@cop.ufl.edu.

<sup>&</sup>lt;sup>†</sup> University of Florida.

<sup>\*</sup> University of Guam.



of the double bond could not be readily deduced from NMR data in CDCl3 because of complete overlap of <sup>1</sup>H NMR signals for H-34 and H-35 for the major conformer; yet the signal for H-34 of the minor conformer ( $\delta_{\rm H}$  6.54) appeared as a doublet with a 15.6 Hz coupling, suggestive of E configuration. To confirm, NMR data were recorded in benzene- $d_6$  (Table S1, see Supporting Information); the conformational ratio changed to 7:3, and signals for H-34 and H-35 were separated for both conformers. The coupling constant between both protons could then be measured in signals of both conformers (15.7 Hz), proving the hypothesized E configuration. The sequence between the spin systems leading to the cyclic structure for 1 was established by HMBC analysis in CDCl<sub>3</sub> and benzene- $d_6$  (Table 1 and Table S1). Furthermore, strong ROESY correlations between H-28a of the moCys unit ( $\delta$  2.76/2.52 in CDCl<sub>3</sub>) and the NH of the *O*-Me-Tyr residue ( $\delta$  5.89/6.18 in CDCl<sub>3</sub>) for both conformers confirmed the assignments.

Close similarity of <sup>1</sup>H and <sup>13</sup>C NMR shifts and  $J_{H,H}$  values of **1** and with *E*-dehydroapratoxin A (**5**) were indicative of the same relative configuration in the polyketide chain. Furthermore, both compounds exist as conformers (3:2) in CDCl<sub>3</sub> (Table 1) due to restricted rotation around the *O*-Me-Tyr-*N*-Me-Ala amide bond.<sup>4</sup> However, in contrast to compound **5**, the *trans* isomer was preferred in apratoxin E (**1**) since a ROESY cross-peak between the  $\alpha$ -protons H-14 and H-18 was observed for the minor (*cis*) but not the major conformer. 1D TOCSY experiments assisted in deducing the coupling constants of those complex multiplets that otherwise were obscured due to extensive signal overlap for both conformers (Table 1).

The absolute configurations of all the amino acid units in the peptide portion of 1 were established as L by chiral HPLC analysis

of the acid hydrolysate. The configuration at C30 is speculated to be *S* based on biosynthetic grounds, since the other natural apratoxins from Apra Harbor also occur with 30*S* configuration.<sup>3,4</sup> The potent cytotoxicity of **1** (see below) furthermore required a 37*S* configuration since recent structure—activity relationship studies for apratoxin analogues have shown that *S* configuration at C37 is essential for activity and inversion would lead to complete loss of activity.<sup>7</sup> Consequently, and consistent with biosynthetic considerations as well, the proposed absolute configuration is 37*S*,39*S*, as in the other apratoxins from the same collection site.<sup>3,4</sup>

Apratoxin E (1) was tested for antiproliferative activity against several human cancer cell lines including HT29 colon adenocarcinoma, HeLa cervical carcinoma, and U2OS osteosarcoma cells, alongside with a ratoxin A (2) and E-dehydroapratoxin A (5) as well as other natural products with nanomolar activity (actinomycin D, paclitaxel) for comparison. Compound 1 displayed lower IC<sub>50</sub> values than its closest analogue, compound 5 (Table 2), yet was 5to 15-fold less active than apratoxin A (2). The reduced activity compared with apratoxin A (2) was expected since the effect of dehydration in the long-chain polyketide unit (2 vs 5) was previously established.<sup>4</sup> Interestingly, the cytotoxicity appears to correlate with the proportion of the trans amide isomer [in NMR solvent CDCl<sub>3</sub>, apratoxins A (2) and C (4), >90%; apratoxin B (3),  $\sim$ 75%; apratoxin E (1),  $\sim$ 60%; *E*-dehydroapratoxin A (5),  $\sim 40\%$ ], which may or may not have biological relevance, as the aqueous solution structure could be different.<sup>4</sup> Furthermore, the retention of significant activity despite modification of the moCys unit unique to apratoxin E(1) indicates that the conjugated carbonyl system is not required for activity, which excludes a mechanism of action that involves a Michael addition on this residue. The discovery of this new analogue provides an opportunity to expand the structure-activity relationship in this series of compounds derived both naturally and produced synthetically.<sup>4,7</sup>

As in the case of the apratoxins, cyanobacteria oftentimes produce more than one member of a certain structural class, which is partially due to relaxed specificity of biosynthetic enzymes that can be exploited for precursor-directed biosynthesis.<sup>11</sup> Particularly, polyketide synthases of prokaryotes can show considerable substrate tolerance,<sup>12</sup> but also substrate binding pockets of nonribosomal peptide synthetase adenylation domains may accept different amino acids within a group of polar or nonpolar amino acids,<sup>13</sup> a phenomenon that commonly leads to replacement of valine with isoleucine in a given cyanobacterial structure, for example.<sup>14</sup> Natural structural diversification may also be due to slight genetic variation or a slightly changed environment. For example, comparison of apratoxin A (2) and E (1) structures suggests that two C-methylation domains are inactivated in the biosynthetic gene cluster from the apratoxin E(1) producer or that acetate is incorporated during chain elongation instead of propionate. Conversely, dehydratase and enoyl reductase activity are putatively responsible for the observed unsaturation in the polyketide section of 1, and an additional enoyl reductase, which then appears only functional in the apratoxin E (1) producer, could account for the saturation of the moCys unit;<sup>15</sup> however, this remains to be shown in biosynthetic studies. Many filamentous cyanobacteria including L. bouillonii also grow in patches, and each patch may be different genetically, or at least chemically, as a result of differential gene transcription in response to different environmental stimuli. Even within one homogeneousappearing collection made in a confined area of Apra Harbor, we found heterogeneity based on 16S rDNA sequences.<sup>4</sup> Some of these circumstances may be the reason that we have discovered apratoxin E (1) and encountered only traces of apratoxin A (2) from this collection; we had found the latter in virtually every sample of this organism obtained since 1991 at this site,<sup>3,16</sup> however, at slightly shallower depths.

$\delta_{\rm H} (I \text{ in Hz}) \qquad \delta_{\rm C},  {\rm m}$
0, dd (7.8, 7.8) 559.3, d 5. m 29.5, t
3, m 5, m 25.3, t
5, m 7, m 47.4, t 8, m
5, d (11.4) 56.6, d 56.6, d
0, m 33.4, d 33.4, d 25.6, t
5, m 3, t (7.2) 10.07, q 0, d (6.6) 13.8, q
1, s 31.5, q 1699. s
8, q (6.6) 60.5, d 4, d (6.6) 14.1, q 4, s 36.6, q
0, ddd $(10.5, 9.6, 4.8)$ $170.2$ , s $5$ , dd $(-12.2, 10.5)$ $37.9, t$ $37.9, t$ $37.9, t$
5, d (9.0) 128.3, s 3, d (9.0) 130.5, d 1, d (9.0) 113.8, d
82, s 55.2, q 55.2, q 9.6)
5, m 32.5, t 32.6, t
1, m 29.9, t 7
75.9, d 1, dd (-10.6, 7.7) 37.6, t 4, dd (-10.6, 2.0)
5, m 166.2, s 126.5, d
4, m 144.2, d 144.2, d 37.36, t 37.36, t
3, dddd (-13.1, 11.9, 3.7) 3, ddddd (11.9, 11.9, 6.6, 4.3, 2.4) 29.4, d 9, ddd (-14.3, 12.0, 2.4) 38.4, t
9, ddd (-14.3, 11.9, 1.7) 9, dd (12.0, 1.7) 24, 6, d
: 0.876, s 3 × 26.0
5. d (6.6) 20.6. a

Table 1. NMR Data for Both Conformers of Apratoxin E (1) in CDCl<sub>3</sub> (ratio 3.2) at 600 MHz (<sup>1</sup>H) and 150 MHz (<sup>13</sup>C)<sup>a</sup>

Table 2. Antiproliferative Activity Data for Various Apratoxins and Other Antineoplastic Natural Products

	IC <sub>50</sub> in nM (HT29)	IC <sub>50</sub> in nM (HeLa)	IC <sub>50</sub> in nM (U2OS)
apratoxin E (1)	21	72	59
apratoxin A (2)	1.4	10	10
E-dehydroapratoxin A (5)	41	121	160
actinomycin D	2.9	0.4	3.2
paclitaxel	1.9	1.5	24

## **Experimental Section**

General Experimental Procedures. Optical rotations were measured on a Perkin-Elmer 341 polarimeter. IR data were recorded on a Bruker Vector 22 IR spectrometer. UV spectra were recorded on a SpectraMax M5 (Molecular Devices). <sup>1</sup>H, <sup>13</sup>C, and 2D NMR spectra were recorded in CDCl<sub>3</sub> and benzene-d<sub>6</sub> on a Bruker Avance II 600 MHz spectrometer equipped with a 1 mm triple-resonance high-temperature superconducting cryogenic probe using residual solvent signals ( $\delta_{\rm H}$  7.26,  $\delta_{\rm C}$ 77.0 for chloroform;  $\delta_{\rm H}$  7.16,  $\delta_{\rm C}$  128.0 for benzene) as internal standards. HSQC and HMBC experiments were optimized for  ${}^{1}J_{CH} =$ 145 and  $^{\it n}J_{\rm CH}=$  7 Hz, respectively. Mixing times of 60 and 500 ms, respectively, were used to acquire TOCSY and ROESY data. HRMS data was obtained using an Agilent LC-TOF mass spectrometer equipped with an APCI/ESI multimode ion source detector.

Biological Material. Lyngbya bouillonii strain PS372 was collected by snorkeling from 12 to 14 m depth at Finger's Reef, Apra Harbor, Guam, during June 2006 and May 2007. Strain PS372 was found only deeper (12-14 m depth) under overhangs or other partially shaded reef structures, unlike the previously described L. bouillonii from shallow reef parts, which was growing between the branches of scleractinian corals such as Porites sp. and Acropora sp. Strain PS372 was mainly found overgrowing the green alga Halimeda sp. and occasionally crustose coralline algae. Unlike L. bouillonii from shallow parts of Finger's Reef (1-3 m), no shrimps were found to be associated with the deep strain PS372. Microscopic evaluation of the voucher material confirmed that size and morphology of the filaments and cells were consistent with L. bouillonii described by Hoffmann and Demoulin in 1991.<sup>17</sup> Cell size of L. bouillonii strain PS372 ranged from 6 to 13  $\mu$ m length and 20 to 24  $\mu$ m width with a cell length-to-width ratio ranging from 0.27 to 0.53. A voucher specimen is retained at the University of Guam Marine Laboratory.

Extraction and Isolation. The freeze-dried organism was extracted twice with EtOAc-MeOH (1:1) to afford 6.92 and 1.89 g of extracts, respectively. The combined extract was subjected to flash chromatography over silica gel, eluting with CH<sub>2</sub>Cl<sub>2</sub>, followed by increasing amounts of *i*-PrOH in CH<sub>2</sub>Cl<sub>2</sub>, and finally with MeOH. The mixtures that eluted with 6% i-PrOH (122 mg) and 8% i-PrOH (88 mg) were further subjected to semipreparative reversed-phase HPLC (Phenomenex Ultracarb, ODS 250  $\times$  10 mm, 5  $\mu$ m, 3.0 mL/min; UV detection at 220, 240 nm) using an isocratic system of 80% aqueous MeCN for 30 min; 80-100% MeCN for 30-40 min; and 100% MeCN for 40-60 min to afford apratoxin E (1) (5.1 mg,  $t_{\rm R}$  34.0 min).

Apratoxin E (1): colorless, amorphous solid;  $[\alpha]^{20} = -69$  (c 0.12, MeOH); UV (MeOH)  $\lambda_{max}$  (log  $\epsilon$ ) 220 (4.28), 270 (3.34) nm; IR (film) v<sub>max</sub> 2956, 2923, 2868, 1740, 1642, 1441, 1247, 1175 cm<sup>-1</sup>; <sup>1</sup>H NMR, <sup>13</sup>C NMR, COSY, and HMBC data, see Table 1 (in CDCl<sub>3</sub>) and Table S1 (benzene- $d_6$ ); HRESI/APCIMS m/z [M + H]<sup>+</sup> 796.4699 (calcd for C<sub>43</sub>H<sub>66</sub>N<sub>5</sub>O<sub>7</sub>S, 796.4683).

Acid Hydrolysis and Chiral HPLC Amino Acid Analysis. A sample of compound 1 (0.3 mg) was treated with 6 N HCl (0.5 mL) at 110 °C for 18 h. The hydrolysate was concentrated to dryness, resuspended in H<sub>2</sub>O (100  $\mu$ L), filtered, and analyzed by chiral HPLC for amino acids [column, Phenomenex Chirex phase 3126 (D) ( $4.6 \times 250$  mm); solvent, 2 mM CuSO<sub>4</sub>-MeCN (95:5 or 90:10); detection at 254 nm]. The retention times ( $t_R$ , min) of the amino acids in the hydrolysate of apratoxin E (1) matched those of N-Me-L-Ala (8.4), L-Pro (10.5), N-Me-L-Ile (25.4), but Notes

not those of N-Me-D-Ala (8.8), D-Pro (21.0), N-Me-L-allo-Ile (24.0), N-Me-D-Ile (34.5), N-Me-D-allo-Ile (34.6) (solvent 95:5, flow rate 0.75 mL/min). The HPLC profile of the hydrolysate (solvent 90:10, flow rate 1.0 mL/ min) also showed additional peaks corresponding to O-Me-L-Tyr and L-Tyr (arising from O-demethylation under the reaction conditions) at 105 and 26.2 min, respectively, while a peak for D-Tyr (27.8) was not detected in the hydrolysate.

Cell Culture. Cell culture medium was purchased from Invitrogen and fetal bovine serum (FBS) from Hyclone. Cells were propagated and maintained in DMEM medium (high glucose) supplemented with 10% FBS at 37 °C humidified air and 5% CO2.

Cell Viability Assays. Cells were plated in 96-well plates (HeLa, 3000 cells; U2OS, 5000 cells; HT29, 10 000 cells) and 24 h later treated with various concentrations of compound 1, other apratoxins, actinomycin D, paclitaxel, or solvent control (1% EtOH for apratoxins or 1% DMSO for other drugs). After 48 h of incubation, cell viability was measured using MTT according to the manufacturer's instructions (Promega).

Acknowledgment. Funding was provided by the James & Esther King Biomedical Research Program, Grant No. 06-NIR07 (to H.L.). The authors gratefully acknowledge the NSF for funding through the External User Program of the National High Magnetic Field Laboratory, which supported our NMR studies at the Advanced Magnetic Resonance Imaging and Spectroscopy (AMRIS) facility in the McKnight Brain Institute of the University of Florida. We thank J. R. Rocca for assistance in acquiring NMR data. The 600 MHz 1 mm triple-resonance HTS cryogenic probe was developed through collaboration between the University of Florida, NHMFL, and Bruker Biospin. Mass spectral analyses were performed at the UCR Mass Spectrometry Facility, Department of Chemistry, University of California at Riverside. P.J.S. acknowledges NIH MBRS SCORE grant SO6-GM-44796-16a for support. This is University of Guam Marine Laboratory contribution number 614.

Supporting Information Available: 1D and 2D NMR spectra in  $CDCl_3$  and in benzene- $d_6$ ; NMR data in benzene- $d_6$  (Table S1). This material is available free of charge via the Internet at http://pubs.acs.org.

## **References and Notes**

- (1) Tan, L. T. Phytochemistry 2007, 68, 954-979.
- (2) Luesch, H.; Chanda, S. K.; Raya, R. M.; DeJesus, P. D.; Orth, A. P.; Walker, J. R.; Izpisúa Belmonte, J. C.; Schultz, P. G. Nat. Chem. Biol. 2006, 2, 158-167.
- (3) Luesch, H.; Yoshida, W. Y.; Moore, R. E.; Paul, V. J.; Corbett, T. H. J. Am. Chem. Soc. 2001, 123, 5418-5423.
- (4) Luesch, H.; Yoshida, W. Y.; Moore, R. E.; Paul, V. J. Bioorg. Med. Chem. 2002, 10, 1973-1978.
- Gerwick, W. H., presented at the US-Japan Symposium on Marine Natural Products, Park City, UT, July 2007.
- (6) Chen, J.; Forsyth, C. J. J. Am. Chem. Soc. 2003, 125, 8734-8735.
- (7) Ma, D.; Zou, B.; Cai, G.; Hu, X.; Liu, J. O. Chem.-Eur. J. 2006, 12, 7615-7626.
- (8) Doi, T.; Numajiri, Y.; Munakata, A.; Takahashi, T. Org. Lett. 2006, 8 531-534
- Luesch, H.; Yoshida, W. Y.; Moore, R. E.; Paul, V. J. J. Nat. Prod. 1999, 62, 1702-1706
- (10) The numbering system for compounds 2-5 was adapted to allow direct comparison of NMR data.
- (11) Magarvey, N. A.; Beck, Z. Q.; Golakoti, T.; Ding, Y.; Huber, U.; Hemscheidt, T. K.; Abelson, D.; Moore, R. E.; Sherman, D. H. ACS Chem. Biol. 2006, 1, 766-779.
- (12) For a recent example, see: Watts, R. E.; Tse, M. L.; Khosla, C. Biochemistry 2007, 46, 3385-3393.
- (13) Challis, G. L.; Ravel, J.; Townsend, C. A. Chem. Biol. 2000, 7, 211-224. (14) Hoffmann, D.; Hevel, J. M.; Moore, R. E.; Moore, B. S. Gene 2003,
- 311, 171-180. (15) For a review on polyketide biosynthesis, see, e.g.: Hopwood, D. A.
- Chem. Rev. 1997, 97, 2465-2497, or other reports in the same issue. (16) Luesch, H. Ph.D. Thesis, University of Hawaii, 2002.
- (17) Hoffmann, L.; Demoulin, V. Belg. J. Bot. 1991, 124, 82-88.

NP700717S